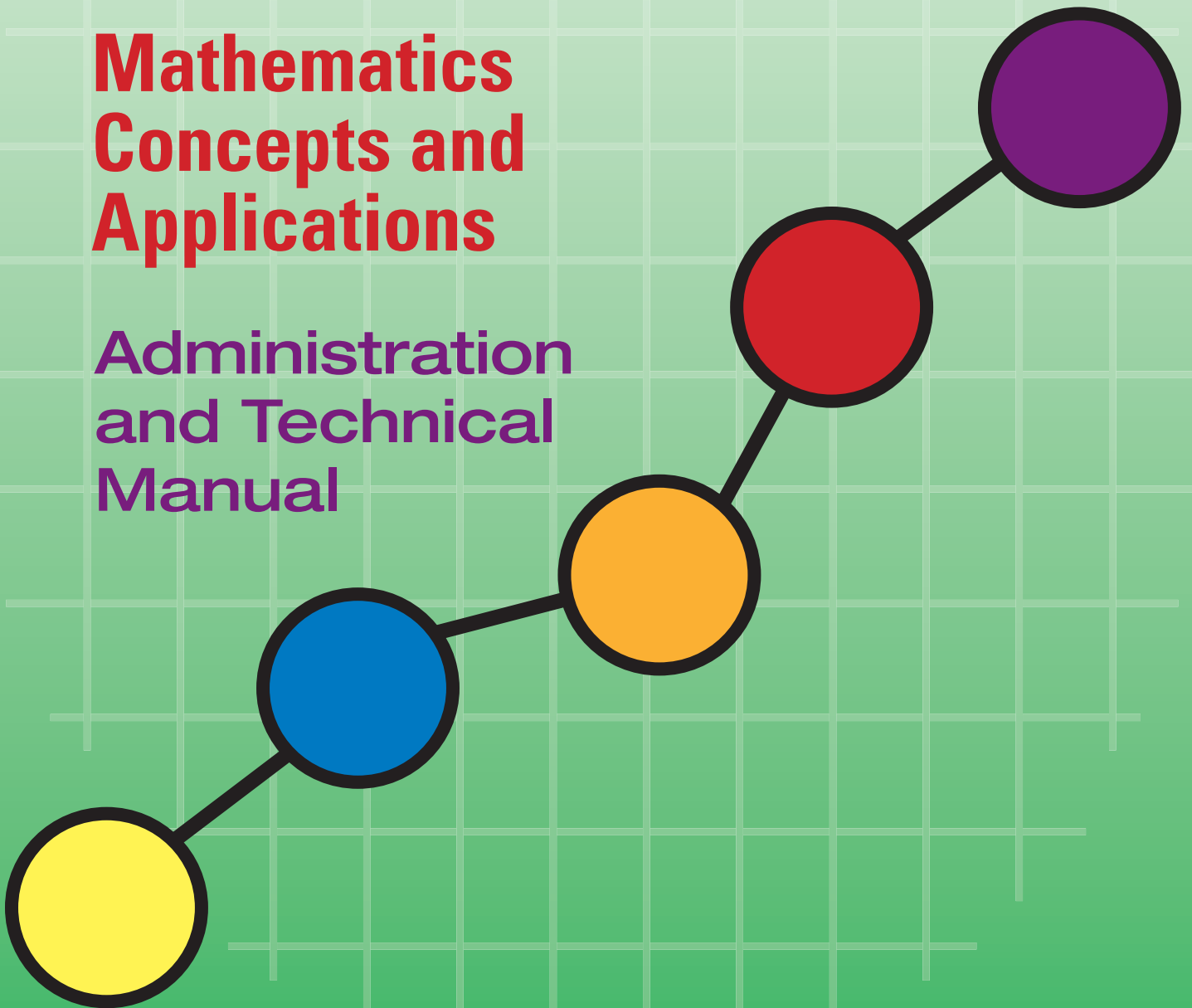




AIMSweb[®]

Mathematics Concepts and Applications

Administration
and Technical
Manual



P.O. Box 599700 San Antonio, TX 78259 800.627.7271 www.PsychCorp.com

Copyright © 2009 NCS Pearson, Inc. All rights reserved.

Warning: No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the copyright owner.

Pearson, the **PSI logo**, **PsychCorp**, and **AIMSweb** are trademarks in the U.S. and/or other countries of Pearson Education, Inc., or its affiliate(s).

Portions of this work were previously published.

Printed in the United States of America.

PEARSON




Table of Contents

Acknowledgements	v
----------------------------	---

Preface	vii
-------------------	-----

Section 1

Introduction	1
Overview	1
Mathematics in a CBM Context	1
Using This Manual	4
User Qualifications, Test Security, and Copyright Restrictions.	4

Section 2

Administration and Scoring	5
Administration Guidelines	5
Before You Test	5
Administration Directions	7
Scoring Guidelines.	9
After Testing.	9
Using M-CAP.	12

Section 3

Development of the AIMSweb M-CAP	17
Test Blueprint	17
Item Development	18
Bias and Content Validity Review	18
Test Construction	18
Major Research Stages	19
Pilot Stage	19
National Field Testing Stage	20
Finalization and Probe Selection Stage	21
Scoring	21
Customer Input.	21
M-CAP Scoring: The New Process	22

Tables

Table 1.1 NCTM Curriculum Focal Points by Grade 2

Table 3.1 Domains Evaluated by M–CAP by Grade 17

Table 3.2 Assessment Lab Anchor Probes Item Count by Grade 19

Table 3.3 Administration Time Limit by Grade 20

Table 3.4 National Field Testing Item and Probe Count by Grade 20

Figures

Figure 2.1 Example of a Scored Answer Key 9

Figure 2.2 Examples of Acceptable Answer Variants 11

Figure 2.3 Example of an M–CAP Individual Student Report 12

Figure 2.4 Example of an M–CAP Progress Monitoring Improvement Report 14

Figure 3.1 Example of Scoring Key Mark-Up 22

Figure 3.2 Example of a Multiple-Part Item: Grade 7 23

Figure 3.3 Example of a Multiple-Part Item: Grade 8 23

Appendixes

Appendix A Technical Adequacy and Data Tables. 25

Appendix B Administration Directions for Grades 2–6 and Grades 7–8 29

Appendix C Further Scoring Guidance 33

References 53



Acknowledgements

As with any complex project, myriad professionals and teams were involved throughout the course of developing the AIMSweb *Mathematics Concepts and Applications*. Although not everyone who touched this product can be named, the hard work and dedication of each of those individuals was important and is greatly appreciated.

Several outside contractors were involved in the early development and review of the AIMSweb *Mathematics Concepts and Applications* (M–CAP) content. The efforts of the Manya Group, led by Pooja Saxena, in working with the content development team on the initial anchor probes for each grade, as well as the subsequent equivalent probes, were integral to the success of this project. The team also thanks the following professional educators for their excellent work in reviewing the anchor probes: Anne Moss, Joseph Lucas, Kerri Shields, Kerry Bollman, Mike Roubison, Shanna Steed, Sue Stiefel, and Tammy Grimes. Their feedback and insight, based on their daily work in the schools and familiarity with local curricula, were invaluable in crafting and validating stable foundational anchors for each grade from which to develop strong equivalent probes. Brandon Anderson’s detailed working of each problem to ensure the accuracy of each item response was an enormous contribution. The expertise of Dave Mellot and DKM Productions was key in refining the format and continuity of the M–CAP probes. We are grateful to Dr. Mark Shinn for his input and consultation during this project.

For their work in proposing, nurturing, and championing this latest addition to the AIMSweb system of assessment, much appreciation goes to Dr. Tom Cayton, Senior Director of Clinical Test Development, Clinical Assessment; Dr. Lawrence Weiss, Vice President of Test Development, Clinical Assessment; Gene Bowles, Vice President of Digital Products and Solutions, Clinical Assessment; Carol Watson, President and General Manager, Clinical Assessment; and Dr. Aurelio Prifitera, Group President and CEO of Clinical Assessment/Worldwide.

Many individuals from different departments at Pearson Assessment and Information, were integral to the development of M–CAP. The development team is very grateful for the contributions of the following individuals: Lisa Yang, Mark Cooley, Ariane Zamot, Kurt Johnson, Jo-Ann Cuellar, Leanne Smith, Ruth Mendez, Elizabeth Rodriguez, Priscilla Villanueva, Carlos Ramirez, John Ramirez, Colleen McAndrew, Rene Martinez, Martha Chichil, Darrell Stiffey, Sheri Schafhauser, Dave Howe, Victoria Locke, Maurya Buchanan, James Henke, David Houghton, Jay Anderson, Julie Page, Cher Whitbeck, Tania Maltos, David Quintero, Glen Larson, Terri Davis, Carol Hansen, Mary Kay Markfelder, Uros Martinovic, Jessica Mehle, Deanna Miller, Andrea Olson, Scott Overgaard, Heather McLaughlin, Lora Oberle, John Bielinski, Jayme Lyon, Lauren Gray, Stephanie Tong, Erik Gallemore, Alanna Carmichael, Anne Trominski, Weslea Miller, Ysela Abastta, Joe Libby, Dawn Dunleavy, Dr. Mark Daniels, and Dr. Jianjun Zhu.

Finally, the dedication and tireless efforts of the AIMSweb *Mathematics Concepts and Applications* core development team must be recognized: Mylene Bagalawis, Clinical Assessment Developer, and Michelle Girard Rolfhus, Senior Clinical Assessment Developer, performed myriad tasks and worked with both in-house and outside contributors throughout the life of this project. Both Mylene and Michelle worked countless hours on every component to ensure the successful and timely completion of the project. Neal Campbell, Project Manager, skillfully coordinated the many moving parts of the project and kept the team focused on the goal; Jeanne Kincaid, Production Manager, and Margaret Young Cooley, Editor, worked together with the content team and DKM Productions to ensure that the look and feel of the final product would meet the needs of our customers in ease of use and accessibility of presentation. Hector Solis, Senior Supervisor, Field Research; Marshall Dahl, Director of Data Management; and Dr. Xiaobin Zhou, Psychometrician, gathered, processed, and analyzed a staggering amount of data in an exceedingly brief window of opportunity.

Most importantly, we thank the thousands of children who participated in field testing. Without them, this project most certainly would not have come to fruition. Last, and certainly not least, we would like to thank not only the educators who worked on this project, but all educators working tirelessly every day on behalf of our children. Words alone cannot convey our appreciation for the work each of you do.

Kurt Hulett, Ed.D.
Senior Research Director, AIMSweb



Preface

Mathematics Concepts and Applications (M–CAP) is the newest addition to the AIMSweb system. AIMSweb is an assessment and web-based data management system designed to efficiently screen, benchmark, and progress monitor basic academic skills in reading, writing, and mathematics for students in grades K–8. The system is built around Curriculum-Based Measurement (CBM), an assessment procedure that uses short standard probes of grade-level outcomes of specific curricula. AIMSweb is well suited for RTI implementation and tiered instruction, and better enables schools to monitor the progress of all students and identify those who face challenges in specific areas.

The first CBM mathematics measures were graded probes that included a broad range of computational problems, sampling from within-grade problem types from grades 1–6 and across grade-specific, single-skill facts probes (e.g., subtraction, division; Tindal, Germann, & Deno, 1983). By the late 1980s, the first set of CBM mathematics application probes that included a range of concepts and application problem were developed (Fuchs & Fuchs, 1987; Fuchs, Fuchs, Hamlett, & Stecker, 1991). In their 2008 publication, Fuchs, Fuchs, and Zumeta discuss two possible methods for designing a CBM system. The second method is utilized by AIMSweb and is described as an approach that “systematically sample(s) the year-long curriculum so that each skill is represented and receives the same emphasis on each alternate form” (2008, p. 225). Using this approach, M–CAP provides a research-based approach for screening, benchmarking, and progress monitoring student performance in mathematics for grades 2–8. For each grade, M–CAP provides benchmark and progress monitor probes; all probes have been field tested to ensure each probe is equivalent.

This Manual describes the administration and scoring procedures, test development, and technical information for AIMSweb *Mathematics Concepts and Applications*.

Overview

The AIMSweb *Mathematics Concepts and Applications* (M–CAP) is a test of short duration (8–10 minutes) that assesses the general mathematics problem-solving skills expected in grades 2–8. The test may be administered in a large or small group setting or to individual students.

M–CAP can be used by teachers and other school professionals to quickly screen and monitor mathematics progress. The mathematics domains assessed include number sense, operations, patterns and relationships, data and probability, measurement, data and statistics, geometry, and algebra. These areas are consistent with the 2006 recommendations of the National Council of Teachers of Mathematics (NCTM) regarding curriculum focal points for grades PreK–8 (NCTM, 2006).

For each grade from 2–8, M–CAP provides 33 equivalent probes that can be used for:

- *Benchmark Assessment:* Universal or individual screening to identify students at risk for difficulties in mathematics, as well as classroom or grade-based benchmarking three times per year to track *all* students' performance throughout the school year (e.g., Tier 1).
- *Strategic Monitoring:* More frequent (once or twice monthly) progress monitoring throughout the school year of at-risk students (e.g., Tier 2, Title I). Examiners typically use one M–CAP probe for each strategic monitoring administration.
- *Frequent Progress Monitoring:* Individual goals are progress monitored one to two times per week throughout the year for students with significant performance discrepancies (e.g., Tier 3) or special education individual education programs (IEPs).
- *Special Education Decision Making:* Includes entitlement through RTI, using M–CAP to determine the severity of educational need and individual student benefit from intervention, progress toward IEP goals, and as part of annual and three-year evaluations.
- *Program Evaluation:* Evaluating the efficacy of intervention programs for groups of students, including but not limited to Tier 2 and Tier 3, and the creation of adequate yearly progress (AYP) targets for state-based mathematics tests or other criterion measures.

Mathematics in a CBM Context

According to Fuchs, Fuchs, and Zumeta (2008), “Historically, mathematics assessments, particularly those for students with learning disabilities, have focused on computation skills despite that students also need to learn to reason mathematically and apply computation skills to a variety of situations” (p. 226). As a result of the demand for students to demonstrate mathematical knowledge

and skills beyond basic computation, additional assessment is recommended to test student knowledge of mathematical concepts and applications. The authors continue:

In an effort to address this limitation, CBM for mathematics currently consists of two mathematics domains: CBM computation (CBM–COMP) and CBM–APP. CBM–COMP monitors students’ progress on fundamental computation skills, including single-digit addition through addition, subtraction, multiplication, and division of decimals and fractions. Computation skills are assessed in isolation, without emphasis on how these skills are applied to problem solving settings (Fuchs et al., 1994). Mathematics domains assessed within CBM–APP include counting, number concepts, number naming, measurement, money, grid reading, charts and graphs, fractions, decimals, applied computation, and word problems. CBM–APP is designed to be used throughout the school year to monitor students’ mathematics progress on concepts and applications. (Fuchs, Fuchs, & Zumeta, 2008, p. 226).

National Council of Teachers of Mathematics

Since the late 1980s, the National Council of Teachers of Mathematics (NCTM) has shaped the national perception regarding what is important for students to learn in mathematics by developing standards designed to ensure that students have the ability to use mathematics to solve real-life problems. First published in 1989 and revised in 2000, the NCTM *Curriculum and Evaluation Standards* has set the stage for a continuous discussion among educators of what should be taught in mathematics and when it should be taught. In its 2006 publication *Curriculum Focal Points for Prekindergarten through Grade 8 Mathematics: A Quest for Coherence*, the NCTM provided further clarity regarding mathematics expectations by grade from kindergarten through grade 8. Table 1.1 presents excerpts of the revised NCTM focal points that form the expectations for grades 2 through 8.

Table 1.1 NCTM Curriculum Focal Points by Grade

Grade	Focal Point	M–CAP	M–COMP
2	Number and Operations: Base Ten, Place Value	X	
	Number and Operations and Algebra; Basic Facts		X
	Measurement: Linear Measurement and Length	X	
3	Number and Operations and Algebra: Understanding of Facts		X
	Number and Operations: Fractions	X	X
	Geometry: Two-Dimensional Shapes	X	
4	Number and Operations and Algebra: Facts and Fluency with Whole Numbers		X
	Number and Operations: Decimals, Fractions and Decimals	X	X
	Measurement: Two-Dimensional Shapes	X	
5	Number and Operations and Algebra: Fluency with Whole Numbers	X	X
	Number and Operations: Decimals, Fractions and Decimals	X	X
	Geometry and Measurement: Three-Dimensional Shapes	X	
6	Number and Operations: Fluency with Multiplication/Division of Fractions and Decimals		X
	Number and Operations: Ratios and Rate	X	
	Algebra: Expressions and Equations	X	
7	Number and Operations and Algebra and Geometry: Proportionality and Similarity	X	
	Measurement and Geometry and Algebra	X	
	Number and Operations and Algebra: Rational Numbers and Linear Equations	X	
8	Algebra: Linear Equations	X	
	Geometry and Measurement: Space, Figures, Angles	X	
	Data Analysis and Number and Operations and Algebra: Data Sets	X	

NCTM, 2006

In 2008, the President’s National Mathematics Advisory Panel (NMAP) released the results of its investigation into what strategies were successful in mathematics instruction. These results, as they relate to the use and purpose of formative evaluation, are excerpted below:

Based on its review of the research, the Panel recommends regular use of formative assessment, particularly for students in the elementary grades. . . . For struggling students, frequent (e.g., weekly or bi-weekly) use of these assessments appears optimal, so that instruction can be adapted based on student progress.
(NMAP, 2008, p. 31)

The design of the AIMSweb M–CAP is uniquely suited to enable schools to follow the NMAP’s recommendation, as it provides probes that can be administered weekly, biweekly, or monthly, depending on the needs of the student. Further, M–CAP probes are simple to administer and score, time efficient, and sensitive to student improvement.

According to the 2001 National Resource Council (NRC) report *Adding It Up*, the mathematics curricula for grades K–8 comprise a number of domains, of which an understanding of number concepts and operations are deemed crucial, and which the authors term *mathematics proficiencies* (p. 5). The authors state:

Mathematical proficiency, as we see it, has five strands:

- conceptual understanding—comprehension of mathematical concepts, operations, and relations
- procedural fluency—skill in carrying out procedures flexibly, accurately, efficiently, and appropriately
- strategic competence—ability to formulate, represent, and solve mathematical problems
- adaptive reasoning—capacity for logical thought, reflection, explanation, and justification
- productive disposition—habitual inclination to see mathematics as sensible, useful, and worthwhile, coupled with a belief in diligence and one’s own efficacy.

The most important observation we make about these five strands is that they are interwoven and interdependent.

(NRC, 2001, p. 5)

The M–CAP probes were designed with this in mind, while also incorporating the mathematical domains identified by the NCTM 2006 standards. Furthermore, M–CAP is designed to reflect the NRC’s recommendations regarding instruction, with a focus on ensuring students’ problem-solving, logical reasoning, and application of analytical skills to problems (NRC, 2001). See Section 3, Table 3.1 for the domains by grade that M–CAP evaluates.

Using This Manual

This Manual provides the M–CAP administration and scoring guidelines, including information regarding the instrument’s development, standardization, and technical attributes. It includes three sections and three appendices.

Section 1 contains introductory information about the M–CAP, explains the purpose of the test, provides a summary of concepts and skill sets measured by M–CAP, and describes test features and the organization of this Manual. Section 2 provides general testing considerations, administration directions, and scoring information. Section 3 provides the development of the test and its standardization. Appendix A presents the technical adequacy and data tables. Appendix B represents the administration directions for grades 2–6 and for grades 7–8. Appendix C provides further scoring guidance for determining the correctness of response variations that may not exactly match the Answer Key.


User Qualifications, Test Security, and Copyright Restrictions

M–CAP was designed for administration by a range of educators, including general and special education teachers, school psychologists, and administrators. Before administering any M–CAP probe, the examiner must read Section 2 of this Manual. Please note that adhering to the administration and scoring rules as established during standardization is essential for accurately obtaining and interpreting results. M–CAP is a standardized test. Any deviation from the administration and scoring procedures described in this Manual may invalidate the results and the normative scores should not be reported or used. Please see Section 2 for further guidance regarding accommodations.

Examiners who use M–CAP are responsible for ensuring that test materials, including the probes, remain *secure*. Individual or collective test items should not be copied or disclosed to anyone but the appropriate school personnel; doing so would compromise the validity of M–CAP and reduce the value of the test as a measurement tool. Under no circumstance should test materials be resold or displayed in locations where unqualified individuals can purchase or view partial or complete portions of M–CAP. This includes personal internet websites and internet auction sites.

All test items, data, and test materials are copyrighted. The Rights and Permissions Department of Pearson Education, Inc., must approve *in writing* the copying or reproduction of any test materials. The only exception to this requirement would be the copying of a completed M–CAP probe for the purpose of conveying a student’s records to another qualified educator.

The aforementioned user qualifications, test security policy, and copyright restrictions are consistent with the guidelines set forth in the *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999).



Section 2

Administration and Scoring

Administration Guidelines

This section provides instructions for the administration and scoring of the M–CAP.

Before You Test

Standard Procedures

As stated previously, M–CAP is a standardized test; to make valid normative (national, state, district, or school) decisions, all directions must be followed carefully. Changes in the presentation of the probes, alterations of the instructions given to students, or the inappropriate use of probes as teaching instruments may invalidate any decisions made or conclusions about student performance that are drawn. Before administering a M–CAP probe, it is important that the examiner become familiar with these administration directions.

Testing Time

Administration time varies depending on grade level. Students in grades 2–6 are allotted 8 minutes to complete a probe. Students in grades 7–8 are allotted 10 minutes to complete a probe (see Appendix B for administration by grade). Giving students more time than is allotted is a violation of standardization and will invalidate the obtained scores.

Students should be encouraged to attempt each item before moving on to the next item. Skipping items should be strongly discouraged. Should the examiner notice a student is skipping ahead without attempting each item, the examiner is to provide the following direction: **Try to work each problem. Do not skip around.**

Test Materials

To administer the probe, you will need the following:

- The standardized administration directions found in this Manual.
- A copy of the appropriate M–CAP probe.
- The probe-specific Answer Key.

Note. When you print a probe, the Answer Key is included. **Remove the Answer Key before replication and retain for your use in scoring.**

- A stopwatch or other accurate timer to monitor administration time.
- Sufficient sharpened pencils.

Testing Students with Special Accommodations

The M–CAP probes, like other CBM measures, are intended to be used for screening, benchmarking, and progress monitoring of *all* students through a benchmark assessment approach, and for more frequent progress monitoring of students at risk (e.g., Tier 2) or with severe achievement discrepancies (e.g., Tier 3). The benchmark assessment process generates a normative score; therefore some constraints are required when special accommodations are considered. In general, any significant changes in the standardized testing process preclude the use of any obtained normative score.

Examples of accommodations that have been proposed by AIMSweb users in the past include (a) increasing the amount of test-taking time, (b) having students practice the test beforehand, or (c) providing feedback during the testing process about whether or not an answer is correct. These accommodations are not satisfactory as they are major changes in the standardized procedure. For some students, like any test, the use of M–CAP probes may be inappropriate because the stimulus demands of the test do not match the sensory response capabilities of a specific student. For example, because M–CAP requires pencil-paper test-taking skills, students with severe motor problems may not be appropriate candidates for M–CAP use. However, for students with mild visual impairments, text enlargement may be a satisfactory accommodation. Although the national sample does include students with disabilities, those students were administered the test in the standardized manner, with no special accommodations.

When M–CAP probes are used for frequent progress monitoring, however, the testing process is more amenable to test accommodations because the student's scores are compared to their *own* scores (i.e., are *individually referenced* instead of norm-referenced) over time. For a student with some motor impairment, increasing test time for progress monitoring purposes (e.g., from 10 minutes to 12 minutes) may be acceptable, providing this increase is kept standard throughout the progress monitoring process. It should be noted that any comparison of these scores to normative scores and interpretation would be inappropriate.

Off-Level Testing

In some circumstances, it is appropriate to administer M–CAP probes from a grade other than the student's actual grade level, in addition to the student's grade-level probes. In the assessment world, this process is called *off-level testing*. For almost 30 years, off-level testing with CBM has taken place within a survey-level assessment (SLA; Shinn, 2008). In an SLA, an individual student is tested on a probe from his or her grade level and then in consecutively lower grades until a normative score is obtained that reflects the student's current level of performance. The normative score is that score where the tested student performs as well as the typical student at that grade. For example, a grade 5 student who performs well below average (e.g., < 10th percentile) on a grade 5 M–CAP benchmark probe would be tested using M–CAP benchmarks from successively lower grade-levels (e.g., grade 4, grade 3) until the student's score is in the average range for that grade level (i.e., between the 25th and 75th percentile).

SLA is done for particular purposes, first and foremost in the process of writing individualized goals for frequent progress monitoring (Shinn, 2002a). Additionally, SLA can provide supplemental information about the severity of an achievement–performance discrepancy. A grade 5 student whose score falls within the average range of grade 2 students on a grade 2 benchmark has a larger performance discrepancy, and a more severe mathematics achievement problem, than a grade 5 student whose performance is average when compared to a grade 4 student on a grade 4 benchmark.

For more information about using an SLA, see the AIMSweb manual *Progress Monitoring Strategies for Writing Individualized Goals in General Curriculum and More Frequent Formative Evaluation* (Shinn, 2002a).

During Testing

Arranging the Test Environment

Getting accurate results depends on how the test environment is arranged. You may administer the M–CAP in full-class groups, in small groups (3–4 students), or individually. In each case, the examiner must carefully monitor student participation and effort. Keep the following in mind:

- Always use the standardized directions. M–CAP probes are *not* to be used as instructional materials or practice worksheets. Although appropriate practice of mathematics skills is a critical component in mathematics success, using these probes as a teaching tool will invalidate the results of the assessment. Instead, you should use those instructional materials that accompany your district's approved mathematics curriculum.
- Monitor students closely during the administration of a probe. Instruct students to try every item and to not skip ahead or simply cross out more difficult items. Skipping or crossing out more difficult items in favor of easier items will not necessarily improve a student's score and it will adversely affect the value of the data you have gathered. If a student is exhibiting either or both of these behaviors, redirect by saying, **Try to work each problem. Do not skip around.**

Establishing and Maintaining Rapport

A positive relationship between the students and the examiner is an essential part of all assessment situations. An accepting, non-threatening tone will promote rapport, which can be developed and maintained in a number of other ways.

- Prior to administering the probe, the examiner should have a thorough understanding of the procedures for administering, timing, and scoring of responses. This will permit the examiner to pay maximum attention to the students during testing.
- Briefly introduce the tasks in grade-appropriate language. Tell the students that some of the tasks may be easy, while others may be more difficult, and that they are not expected to answer all the questions correctly. Tell the students that they are to try all of the items in order, and that skipping around is not permitted.
- If a student asks a question or requests clarification, redirect him or her to the probe and say: **Read the directions again, and work the problem as best you can. If you still do not understand the problem or are unable to work it, you may move on to the next question.**

Administration Directions

The general M–CAP administration directions for all grades are presented in this section. With the exception of a different time limit for grades 7 and 8, these directions are identical across grades. For the examiner's convenience these directions are presented again in Appendix B, with the grades 2–6 directions separate from the grades 7–8 directions.

For appropriate grades, make sure all cell phones, beepers, and watch alarms are turned off.

The instructions are carefully worded with simple, age-appropriate language. What you say to the students is in **bold print**.

Say to the students:

For grades 2–6: We’re going to take an 8-minute math test.

For grades 7–8: We’re going to take a 10-minute math test.

For all grades: Read the problems carefully and work each problem in the order presented. Do not skip around.

If you do not know how to work a problem, mark it with an X and move on. Once you have tried all of the problems in order, you may go back to the beginning of the worksheet and try to complete the problems you marked.

Write the answers to the problems in the blanks. For multiple choice questions, place the letter (A, B, or C) of the correct answer in the blank.

You do not have to show your work, but you may if that is helpful for you in working the problems.

Keep working until you have completed all of the problems or I tell you to stop.

Do you have any questions?

Answer any questions the students may have, hand the students their probes, and say:

Here are your tests.

Put your name, your teacher’s name, and the date on each page in the space provided.

Do not start working until I tell you to begin.

Allow the students time to write their information on the probe.

Begin.

If a student asks a question or requests clarification, redirect him or her to the probe and say:

Read the directions again, and work the problem the best you can.

If you still do not understand the problem or are unable to work it, you may move to the next question.

When the appropriate time has elapsed (8 minutes for grades 2–6, 10 minutes for grades 7–8), say:

Stop and put down your pencil.

If a student(s) continues to work, re-state:

Stop working now and put down your pencil.

At this time, the examiner should collect the probe(s) and proceed to scoring.

Scoring Guidelines

After Testing

Once the students have completed the M-CAP probes, collect and score. Unlike AIMSweb M-CBM, there is no correct digit or partial-credit scoring; the problem either receives full credit or no credit. If any part of an answer is incorrect, whether it is a one-part or multi-part question, the score for that item is 0. For the user's convenience, the Answer Keys are designed for easy scoring. The teacher circles the score value for a correct answer and a 0 for an incorrect answer. The total score value is then summed at the bottom of the page. Figure 2.1 shows a scored Answer Key for grade 2.

Grade 2, Probe 10 Answer Key			
Item No.	Answer	Correct	Incorrect
1.	6	①	0
2.	157, 720	①	0
3.	10	1	①0
4.	32	①	0
5.	47	①	0
6.	6	①	0
7.	+	①	0
8.	24, 12	2	①0
9.	56, 60	①	0
10.	13	①	0
11.	90	①	0
12.	428, 413	①	0
13.	8, 15	①	0
14.	8, 24, 32	②	0
15.	–	①	0
16.	C	③	0
17.	$\frac{2}{25}$	2	①0
18.	5, 9, 1	②	0
19.	1.65	2	①0
20.	less	1	①0
21.	C, F, E	②	0
22.	7, 8, 2	①	0
23.	4	②	0
24.	137	①	0
25.	8	2	①0
26.	1, 15	②	0
27.	5	③	0
28.	$\frac{10}{25}$ or $\frac{2}{5}$	②	0
29.	690	③	0
Total		35	

Figure 2.1 Example of a Scored Answer Key

One of the challenges in scoring is determining what to do if an answer deviates from the provided Answer Key, but may still be correct. This is especially important because in M–CAP there is no partial-credit scoring. The criteria used to decide when variant answers are or are not acceptable are based on best practices and professional judgment. The primary goal is to determine if the answer reflects an understanding of the task presented (e.g., a money task response indicates if the student knows how to properly express monetary amounts).

A number of items on the M–CAP result in responses that can be reduced to a simpler form or request that the student estimate or round the answer. For these items, there may be a range of acceptable responses provided on the Answer Key. Figure 2.2 (a–c) shows examples of each type of item and the possible correct answers.

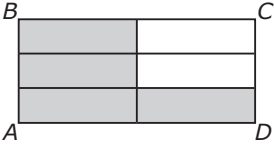
Although alternate acceptable responses are given in the Answer Key for many items, credit may be given for a clearly correct response conveyed in a manner other than the one indicated; this is where the examiner must rely on best practices and professional judgment. To illustrate, grade 7, Item 1 (*Write these numbers in increasing order* [Figure 2.2d]) presents a mix of whole numbers, negative numbers, fractions, and decimals. If the student converts some or all of the options into decimal values, as long as the conversions are correct and the options are correctly placed in order, the item receives full credit.

Appendix C presents examples of the most common variations of correct and incorrect answers seen in the national field-test sample, as well as examples of answers that require the examiner to use his or her judgment in evaluating correctness. Also included are examples of the types of issues that impact scoring decisions, including but not limited to, problems with legibility, reversed numerals, crossed-out responses, and over-correction.

Detailed information on the development of the M–CAP scoring guidelines can be found in Section 3.

A

5



The shaded area is what fraction of rectangle $ABCD$?

$\frac{2}{3}$

Grade 6, Probe 8 Answer Key

Item No.	Answer	Correct	Incorrect
5.	$\frac{4}{6}$ or $\frac{2}{3}$	1	0

B

9

Richard takes out a jelly bean at random from a bag containing 8 blue jelly beans and 1 yellow jelly bean. What is the probability that Richard will take out a blue jelly bean?

8 in 9

Grade 7, Probe 5 Answer Key

Item No.	Answer	Correct	Incorrect
9.	$\frac{8}{9}$ or 8:9 or 8 in 9	1	0

C

22

The number of fairy tale books in a library is 565. The number of science fiction books is 1274. Estimate (to the nearest hundreds place) the total number of both fairy tale books and science fiction books in the library.

1900

Grade 7, Probe 1 Answer Key

Item No.	Answer	Correct	Incorrect
22.	1800 or 1900	2	0

D

1

Write these numbers in increasing order.

$0.7, \frac{-7}{5}, 7.8, \frac{7}{4}, -1.7$

-1.7 -1.4 0.7 1.75 7.8

Grade 7, Probe 2 Answer Key

Item No.	Answer	Correct	Incorrect
1.	$-1.7, \frac{-7}{5}, 0.7, \frac{7}{4}, 7.8$	2	0

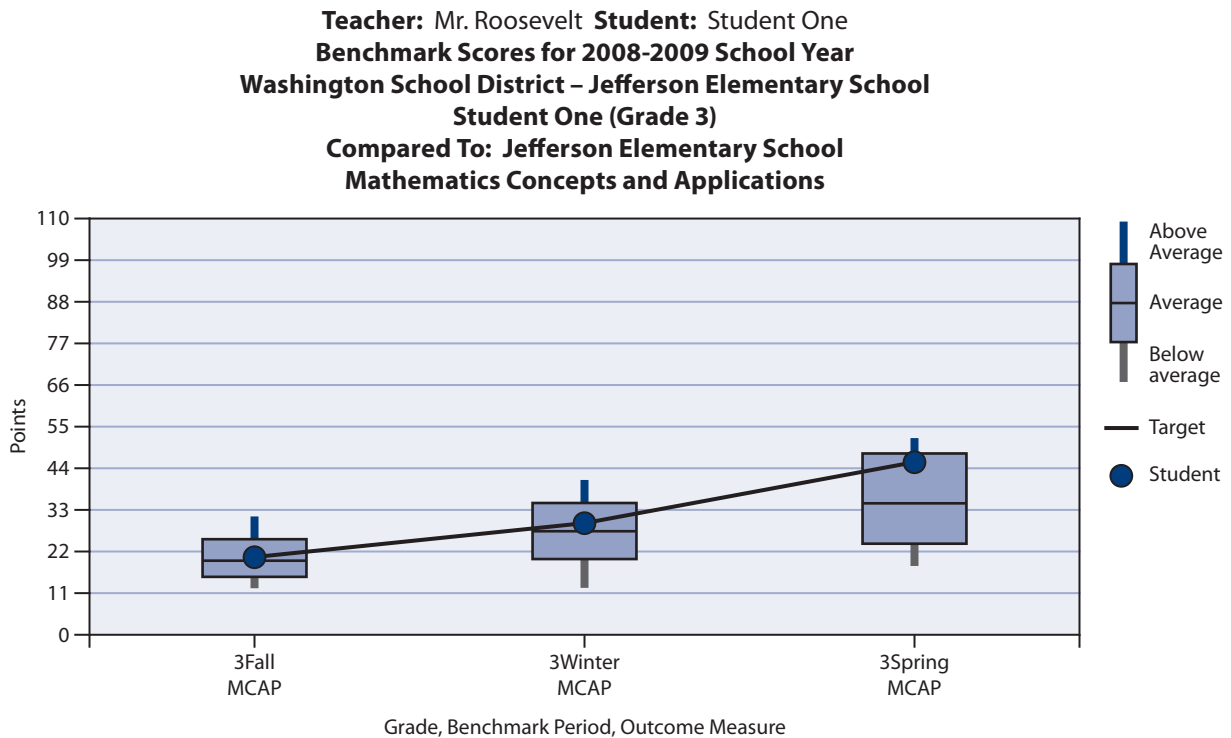
Figure 2.2 Examples of Acceptable Answer Variants

Using M-CAP

Screening and Benchmarking

Using the three probes identified as benchmark probes, all students are tested three times per year; early in the year (e.g., fall) for purposes of screening, and twice more (e.g., winter, spring) for progress monitoring and program evaluation. Students are administered one probe during each benchmarking period.

M-CAP decisions can be made on a normative basis or on a standards-based basis. With a normative decision-making basis, an individual student's scores are compared to the scores of another group of students. With AIMSweb, educators can compare an individual student to a number of different levels of local norms beginning at the classroom and moving to broader contexts of school, school district, state, or a composite of AIMSweb customers.



Copyright © 2009 by NCS Pearson, Inc.

Benchmark Comparison: Jefferson Elementary School

Outcome Measure	Year	Grade	Fall	Winter	Spring	Level of Skill	Instructional Recommendation
Mathematics Concepts and Applications (MCAP)	2008–2009	3	20.0	29.0	45.0	Average	Continue Current Program (Jefferson Elementary School Spring Percentiles)

Student One improved from **20** Points (pts) from Grade 3 Probes at the Fall Benchmark to **45** Points (pts) at the Spring Benchmark. The rate of improvement (ROI) from the Fall Benchmark is **0.7** Points per week. Currently, Student One's score is **Average** compared to Jefferson Elementary School Spring Percentiles. This was a score **at the 64** percentile compared to other students in the Jefferson Elementary School.

Figure 2.3 Example of an M-CAP Individual Student Report

Figure 2.3 presents an example of an M–CAP Individual Student Report. This box-and-whisker chart represents the range of average M–CAP student performance (i.e., scores between the 25th and 75th percentiles). The line extending from the top of the box represents the range of above average M–CAP student performance (i.e., scores between the 75th and 90th percentiles). Scores above this top line represent scores in the upper 10 percent of students in the comparison group. The line extending from the bottom of the box represents the range of below average M–CAP student performance (i.e., scores between the 10th and 25th percentiles). Often, scores at this level are used to identify potential candidates for tiered interventions (e.g., Tier 2). Scores below this bottom line represent scores in the lower 10 percent of students in the comparison group. Typically, scores at this level are used to identify potential candidates for the most intensive of tiered interventions (e.g., Tier 3).

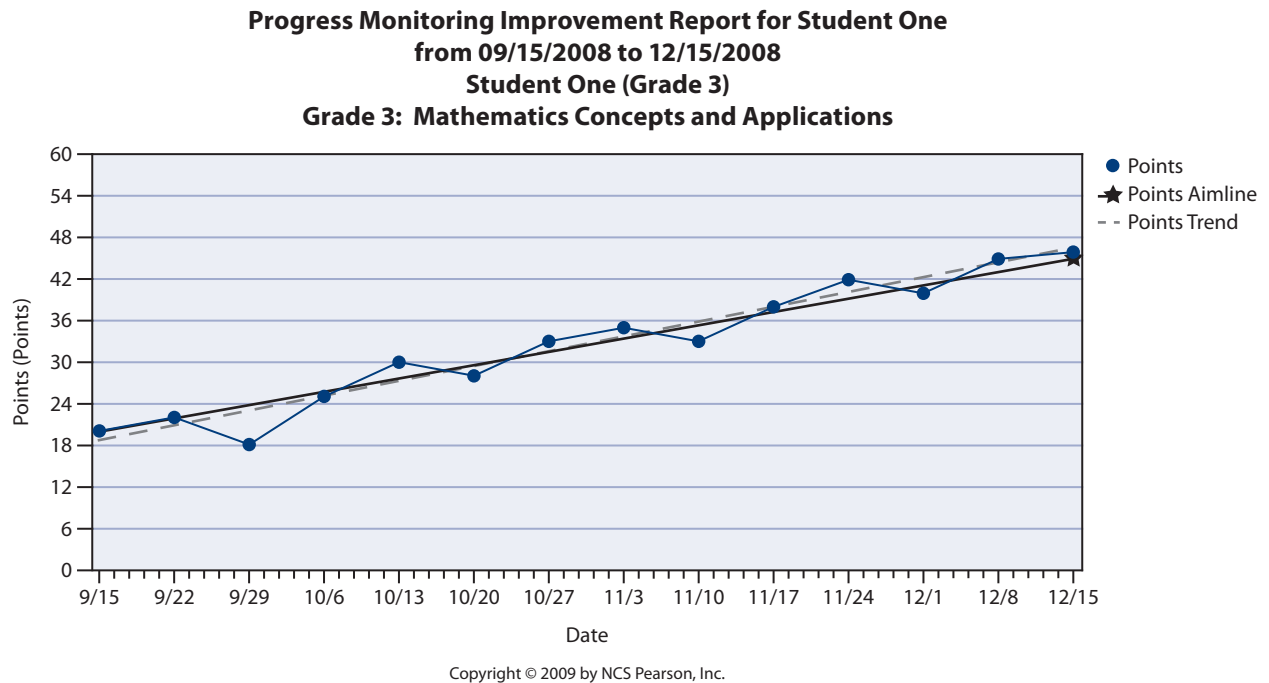
The level of a student’s M–CAP performance over time, and the number of points earned on a particular probe from fall to winter (or from fall to winter to spring), represent that individual student’s rate of improvement (ROI). In a benchmark approach, an ROI score can be obtained for each student tested using the M–CAP. Any individual student’s ROI can be compared to a normative group. Figure 2.3 presents a sample showing a student’s M–CAP performance during the course of a school year relative to a school norm.

When using a standards-based decision-making approach, an individual student’s scores are compared to predicted performance on a high-stakes test (e.g., a state required achievement test). Typically, this predicted performance is a categorical score, such as “Highly Likely to Pass” or “Not Likely to Pass.” To use a standards-based approach for decision making, the relationship of M–CAP to the high-stakes test must be established empirically.

For more detail in how to use M–CAP in a benchmark assessment approach for screening and progress monitoring of all students, see the AIMSweb Training Workbook *Organizing and Implementing a Benchmark Assessment Program* (Shinn, 2002b). This and other manuals can be found on the Downloads tab within your AIMSweb account.

Frequent Progress Monitoring

M–CAP can be used when there is a need to monitor progress on a more frequent basis than that provided by a benchmark assessment approach. This frequency can range from one to two times per month, such as what is observed typically for students receiving Tier 2 interventions, to one to two times per week, which is observed typically for students receiving Tier 3 or special education interventions. In most progress monitoring instances, a single M–CAP probe is used each time. When progress is monitored frequently, an individual student’s ROI, or the “trend line,” is compared to the student’s expected rate of progress, or the “aim line.” The former is calculated mathematically by drawing a line of best fit through the individual data points. The latter is determined through a goal setting process. A sample frequent progress monitoring graph is shown in Figure 2.4.



Goal Statement

In **13.0** weeks, Student One will achieve **45** Points from grade **3** Mathematics Concepts and Applications. The rate of improvement should be **1.92** Points per week. The current average rate of improvement is **2.14** Points per week.

Date	09/15	09/22	09/29	10/06	10/13	10/20	10/27	11/03	11/10	11/17	11/24	12/01	12/08	12/15
Points	20	22	18	25	30	28	33	35	33	38	42	40	45	46
Goal/Trend ROI	1.92/2.14													

Figure 2.4 Example of an M-CAP Progress Monitoring Improvement Report

Goal Setting

AIMSweb M-CAP is a valuable tool for educators to use when setting goals. With national norms unavailable until the fall of 2010, educators will have to use alternatives to setting goals by nationally normed data during the 2009/2010 school year. In this section, these alternative methods for goal setting are discussed. For more detailed direction for setting goals, see *Progress Monitoring Strategies for Writing Individualized Goals in General Curriculum and More Frequent Formative Evaluation* (Shinn, 2002a).

Goal setting for frequent progress monitoring is based on long-standing practices (Fuchs, Fuchs, & Hamlett, 1990; Fuchs, Fuchs, Hamlett, & Whinnery, 1991; Fuchs & Shinn, 1989; Mirkin, et al., 1981; Shinn & Shinn, 2000). An M-CAP goal is written in an individualized yet standard format such as the one below.

In 34 weeks (1 academic year), Betsy will write correct answers to problems earning 40 points on grade 5 M-CAP probes.

First, a time frame is established (e.g., 34 weeks). Frequent progress monitoring goals often are written with two common time frames, (a) an annual focus and (b) an end-of-year focus. An annual time-frame is used when IEP goals are written for students who are receiving special

education. An end-of-year time frame is used when students do not have IEP goals, but are receiving some type of intensive intervention.

Next, the goal material must be identified. Two scenarios are common: First, when an IEP annual goal is written, the goal material is that level of performance (e.g., grade 5 M–CAP probe) that the student is expected to be performing in successfully in 1 year if the special education intervention is successful. Next, the criterion for success must be specified. Educators have four methods for specifying the criterion, each with its own advantages and disadvantages. The preferred method would be a standards-based approach where the individual student’s goal attainment would more likely result in passing a high-stakes test. The other methods are norm-based: Two of the methods for specifying the criterion for success are tied to local or national normative levels of performance. For example, the team may decide to have the individual student answer as many M–CAP problems correctly as would a grade 5 student at the 50th percentile when compared to their local school district. This local norm has the advantage of educators having a familiarity as to what these type of scores “mean” in every day learning. The last method, and the least preferred, is to use a normative ROI. Using a mathematical formula, an average rate of weekly improvement attained from a normative database is multiplied by the time frame to determine the criterion for success.

Section 3

Development of the AIMSweb M-CAP

Test Blueprint

A primary goal in the preliminary stages of development for the AIMSweb *Mathematics Concepts and Applications* (M-CAP) was to determine the content for, and the general proportions of that content at, each grade level (2–8). As each state adheres to different learning standards, developing grade-specific tests that would broadly sample the content taught across the United States was challenging. After analyzing mathematics standards across several states, it was clear that a significant amount of overlap existed; however, no one state could serve as a template for a national assessment.

As a means to determine the most appropriate content to be assessed at each grade level for a national sample, the AIMSweb team researched the most widely accepted and utilized national mathematics standards for grades 2–8. After extensive research, the team determined that the National Council of Teachers of Mathematics *Principles and Standards* (NCTM, 2006) presents the most broadly recognized and accepted standards within the professional community, and therefore is appropriate to serve as the basis for the determination of content at each grade level.

The content coverage on the *Stanford Achievement Test, Tenth Edition* (Stanford 10) serves as a general guideline for determining the proportion of items by learning domains at each grade level. The Stanford 10 was chosen because it is one of the most widely used norm-referenced assessments of mathematics achievement in the United States.

Table 3.1 Domains Evaluated by M-CAP by Grade

	Grade						
	2	3	4	5	6	7	8
Number Sense	✓	✓	✓	✓	✓	✓	✓
Operations	✓	✓	✓	✓	✓	✓	✓
Patterns & Relationships	✓	✓	✓	✓	✓	✓	✓
Measurement	✓	✓	✓	✓	✓	✓	✓
Geometry	✓	✓	✓	✓	✓	✓	✓
Data & Probability	✓	✓	✓	✓			
Algebra				✓	✓	✓	✓
Probability					✓	✓	✓
Data & Statistics					✓	✓	✓

Item Development

Each item was individually developed by professional test developers and experts in the field of mathematics. The development of each item was based on grade-level and domain-specific criteria. Approximately 11,200 items were developed in total. Each item was extensively field tested prior to the final selection of probes for publication (see p. 19 for field testing details). Once the items were completed, the grade-specific probes were constructed based on the Stanford 10 proportions and AIMSweb test construction criteria.

Bias and Content Validity Review

After extensive internal editing, review, and analysis, the probes were evaluated by professional educators (mathematics teachers and content experts) from across the United States. In addition, each anchor probe was reviewed and evaluated by Mark R. Shinn, Ph.D., a regular contributor to the research and practice regarding Curriculum-Based measurement (CBM) and AIMSweb. Reviewers were provided with a rubric and asked to thoroughly evaluate the items and probes. In particular, reviewers were asked to evaluate the items and probes by responding to the following questions:

- Is the item representative of content in your curriculum?
- Is the item grade appropriate?
- Is the item at the appropriate difficulty level?
- Is the vocabulary grade appropriate?
- Are all of the mathematical symbols in the item developmentally appropriate and consistent with your districts approach?
- Is the format of the question appropriate (art, amount of text, easy-to-understand)?
- Does the item need more specific directions?
- Does the potential for bias exist (should any words be removed or replaced?)

In general, the reviewers rated the probes highly and provided positive feedback. The reviewers' feedback on individual items was incorporated into the editorial process for making appropriate adjustments prior to field testing.

Test Construction

An anchor probe was developed for each grade level (2–8). The anchor probe served as the template from which all equivalent probes at that grade were developed. Each anchor probe was constructed by selecting individual items based on multiple criteria. Prior to item selection and placement, all items were field tested and then evaluated based on point biserial correlations, item difficulty, and individual item administration time. Items that did not meet psychometric criteria were removed from the item pool. Item placement on each anchor probe was based on increasing item difficulty. To ensure the sensitivity of the instrument and to maximize the amount of data collected from at-risk learners, easier items were generally placed at the beginning of each probe and more difficult items followed.

In order to maintain proper randomization and representation of item type by domain, items were not placed in exact order of difficulty within each probe. In addition, the placement of multiple item types measuring the same domain in groups of three or more was avoided.

Each equivalent probe was built to replicate the item type proportions, difficulty, and item placement on the anchor probe. For example, Item 1 of the grade 2 anchor probe is a measurement question; therefore Item 1 for each subsequent and equivalent grade 2 probe is the same problem *type* (i.e., measurement) with a different *item* of similar difficulty and construction. Although each item is unique, the intent is to assess the same learning domain with a similar question at the same numbered position on each equivalent probe.

Table 3.2 Assessment Lab Anchor Probes Item Count by Grade

Grade	Item Count by Probe
Grade 2	35
Grade 3	35
Grade 4	35
Grade 5	35
Grade 6	35
Grade 7	35
Grade 8	35

Major Research Stages

Pilot Stage

Three pilot studies were conducted at the beginning of the field research process. The primary goal of the pilot stage was to produce anchor probes that had the desired content coverage and psychometric properties for use in the generation of equivalent probes for national field testing. This stage of development focused on such issues as item content and relevance, adequacy of scale floors and ceilings, appropriate time limits at administration, scoring criteria, and other relevant psychometric properties.

Pilot 1: The anchor probes were administered individually to 56 students in order to capture item-specific data, such as item completion time and clarity of both general and item-specific administration instructions. Individual items that took a disproportionate amount of time to complete were identified and eliminated. In addition, proctors took notes on specific questions that were raised by the students. These qualitative data provided information on how the administration directions were functioning.

Pilot 2: Two probes, the anchor and a “cloned” probe, were group administered to 140 students to determine alternate-form reliability, which was critical for the later development of equivalent probes. In order to determine the time limits for administration at each grade, students were timed and progress was marked at the 6-, 7-, 8-, 9-, 10-, 11-, and 12-minute marks. A rank-order correlation analysis was conducted at each grade level to determine the shortest amount of

time necessary for the test to maintain reliable discriminability. As shown in Table 3.3, the data indicate the following time limits are optimal at each grade.

Table 3.3 Administration Time Limit by Grade

Grade	Time Limit in Minutes
Grade 2	8
Grade 3	8
Grade 4	8
Grade 5	8
Grade 6	8
Grade 7	10
Grade 8	10

Pilot 3: The anchor probes were group administered to another 140 students. The administration was not timed. The intent of this study was to extend the collection of item-specific and probe-level data from Pilot 1 and Pilot 2 to further evaluate the performance of items. Items were examined by multiple criteria, including point-biserial correlation coefficients and *p*-values. In addition, the reliability of the anchor probe was evaluated by split-half correlation and Cronbach's alpha.

National Field Testing Stage

Following the pilot stage, a national field-testing edition of the M-CAP was developed. Data were obtained from a stratified sample of 6,550 students representing some key demographic variables in the population. The national sample was stratified along the following variables: grade, sex, race/ethnicity, socio-economic status, and geographic region. See Appendix A for the demographic information for the national sample.

For national field testing, 45 probes (including the anchor probes) were used. Given the administration time limit, and to avoid the accelerated practice effect from answering questions on the same domain repeatedly, it was determined that only a single set of five probes were to be given to any student. Eleven sets of probes were assembled for each grade. In each set, the anchor probe was always administered first. The remaining four probes were administered in reverse order for approximately half of the sample.

Table 3.4 National Field Testing Item and Probe Count by Grade

Grade	Item Count by Probe	Probe Count by Grade (PMs & BMs)
Grade 2	29	45
Grade 3	29	45
Grade 4	30	45
Grade 5	30	45
Grade 6	29	45
Grade 7	31	45
Grade 8	28	45

Finalization and Probe Selection Stage

Multiple criteria were used to select the most psychometrically-sound equivalent probes. Pearson's product-moment correlation coefficient was examined to assess the consistency of probes within each grade. The average correlation coefficients for each probe with other probes in the set are reported in Table A.6. To evaluate the internal consistency of the probes, Cronbach's alpha and split-half reliability were examined (see Table A.6). Probes that met minimum reliability standards (i.e., average correlation of .80 or greater) were sorted by the mean score. Probes with means that deviated the most from the average were removed. Analysis of the confidence interval at the 95% level using the standard error of measurement (*SEM*) showed that the final selected probes were statistically equivalent to each other in the grade. The aggregated means and *SEMs* were also reported in Table A.6.

Scoring

Customer Input

As an ongoing process, the AIMSweb team solicits feedback from customers regarding the AIMSweb system. In relation to mathematics-specific scoring, customers primarily noted concerns in the following areas: (1) length of time and difficulty of scoring, (2) general presentation and format of test and items which impeded the scoring process, and (3) students skipping around to easier items. The three major customer concerns were taken into account and addressed during the conceptual development of the M-CAP.

To address the length of time and difficulty of scoring issue, the research team evaluated the current process utilized for M-CBM. The critical process (partial-credit scoring) is currently used. Customers noted that counting every numeral is time intensive. The research team was charged with developing a scoring system that would minimize scoring time, maximize sensitivity to growth, control for students who skip to the easiest items, and ensure the psychometric soundness of the process. In addition, because a number of AIMSweb users noted that the test format impeded their ability to score the test quickly (items too close on the page and answers running into other answers), the team was also charged with creating a more user-friendly format.

To address test format, boxes were placed around each item and the spacing between items was increased. A style guide for each grade was developed to ensure consistency of mathematical language, symbols, and overall style. To respond to concerns regarding the impact of students skipping around and completing the easiest items, weights were applied to items based on difficulty (see M-CAP Scoring: The New Process for more detail). Students who have the tendency to skip to the easiest items will not necessarily receive higher scores. In addition, by placing the easiest items at the beginning of each probe, students are less likely to skip ahead.

M-CAP Scoring: The New Process

In M-CAP, answers are either correct or incorrect and each item has a weighted value (1 point, 2 points, or 3 points) based on item difficulty. This new scoring process does not rely on or allow for partial-credit scoring, which significantly decreases the scoring time. The least-difficult items are scored 1 point, items of medium difficulty are scored 2 points, and the most difficult items are scored 3 points. Figure 3.1 illustrates how to use the Answer Keys to score: If the item is correct, the scorer circles the value (1, 2, or 3) in the Correct column. If the answer is incorrect or not answered, the scorer circles the 0 in the Incorrect column. Once each item is scored, the scorer sums the values in the Correct column. The sum is the student's total score for the probe.

Grade 2, Probe 10 Answer Key

Item No.	Answer	Correct	Incorrect
1.	6	①	0
2.	157, 720	①	0
3.	10	1	①
4.	32	①	0
5.	47	①	0
6.	6	①	0
7.	+	①	0
<hr/>			
21.	C, F, E	②	0
22.	7, 8, 2	①	0
23.	4	②	0
24.	137	①	0
25.	8	2	①
26.	1, 15	②	0
27.	5	③	0
28.	$\frac{10}{25}$ or $\frac{2}{5}$	②	0
29.	690	③	0
Total		35	

Figure 3.1 Example of Scoring Key Mark-Up

Regardless of the construction of the item (e.g., some items have multiple parts), each item as a whole is either correct or incorrect—no partial-credit scoring. Even if a student answers only *part* of the item incorrectly, the entire item is incorrect and scored 0. Figure 3.2 shows an example of an item with multiple parts from grade 7. In the example, part 1 is correct but part 2 is incorrect, so the entire item receives a score of 0.

15

Day	Minutes for Paul to Reach the Office
Monday	43
Tuesday	33
Wednesday	35
Thursday	39
Friday	31

On which day did Paul reach the office in the shortest amount of time? Friday

How many minutes more did Paul take to reach the office on Monday than on Thursday? 5

Grade 7, Probe 1 Answer Key

Item No.	Answer	Correct	Incorrect
15.	Friday, 4	1	0

Figure 3.2 Example of a Multiple-Part Item: Grade 7

Figure 3.3 shows an item from grade 8, where the student must fill in 5 blanks. The student filled in 3 out of the 5 blanks correctly and two incorrectly. Although most of the responses are correct, the entire item receives a score of 0. *Note.* The correct order is -3.8, -8/5, 2/7, 2.9, and 3.6.

9

Write these numbers in increasing order.

2.9, $-\frac{8}{5}$, 3.6, $\frac{2}{7}$, -3.8

-3.8 $-\frac{8}{5}$ 2.9 $\frac{2}{7}$ 3.6

Grade 8, Probe 22 Answer Key

Item No.	Answer	Correct	Incorrect
1.	788.008, 788.08, 788.80	1	0
8.	45	1	0
9.	-3.8, $-\frac{8}{5}$, $\frac{2}{7}$, 2.9, 3.6	2	0
10.	80, 135	1	0
11.	136	1	0

Figure 3.3 Example of a Multiple-Part Item: Grade 8



Appendix A

Technical Adequacy and Data Tables

Table A.1 Demographic Characteristics of the Sample, by Grade and Geographic Region

Grade	Geographic Region								Total
	Northeast		Midwest		South		West		
	N	%	N	%	N	%	N	%	
2	85	8.0	305	28.7	584	54.9	90	8.5	1,064
3	71	7.4	230	23.8	564	58.4	100	10.4	965
4	127	12.4	366	35.7	467	45.5	66	6.4	1,026
5	7	0.8	309	35.6	468	54.0	83	9.6	867
6	103	12.0	246	28.7	432	50.3	77	9.0	858
7	48	5.3	269	29.5	464	50.9	131	14.4	912
8	44	5.1	275	32.1	401	46.7	138	16.1	858
Total	485	7.4	2,000	30.5	3,380	51.6	685	10.5	6,550

Note. Row percentages may not sum to 100 due to rounding.

Table A.2 Demographic Characteristics of the Sample, by Grade and Community Type

Grade	Community Type						Total N
	Urban		Suburban		Rural		
	N	%	N	%	N	%	
2	173	16.3	545	51.2	346	32.5	1,064
3	164	17.0	448	46.4	353	36.6	965
4	183	17.8	503	49.0	340	33.1	1,026
5	132	15.2	355	40.9	380	43.8	867
6	151	17.6	324	37.8	383	44.6	858
7	224	24.6	395	43.3	293	32.1	912
8	159	18.5	460	53.6	239	27.9	858
Total	1,186	18.1	3,030	46.3	2,334	35.6	6,550

Note. Row percentages may not sum to 100 due to rounding.

Table A.3 Demographic Characteristics of the Sample, by Grade and Sex

Grade	Sex				Total
	Female		Male		
	N	%	N	%	
2	552	51.9	512	48.1	1,064
3	497	51.5	468	48.5	965
4	506	49.3	520	50.7	1,026
5	430	49.6	437	50.4	867
6	412	48.0	446	52.0	858
7	457	50.1	455	49.9	912
8	445	51.9	413	48.1	858
Total	3,299	50.4	3,251	49.6	6,550

Note. Row percentages may not sum to 100 due to rounding.

Table A.4 Demographic Characteristics of the Sample, by Grade and Race/Ethnicity

Grade	Race/Ethnicity												Total
	African American		American Indian		Asian		Hispanic		White		Other ^a		
	N	%	N	%	N	%	N	%	N	%	N	%	
2	109	10.2	16	1.5	19	1.8	350	32.9	566	53.2	4	0.4	1,064
3	83	8.6	7	0.7	8	0.8	210	21.8	655	67.9	2	0.2	965
4	107	10.4	27	2.6	8	0.8	213	20.8	667	65.0	4	0.4	1,026
5	77	8.9	30	3.5	7	0.8	194	22.4	554	63.9	5	0.6	867
6	45	5.2	23	2.7	19	2.2	271	31.6	499	58.2	1	0.1	858
7	78	8.6	4	0.4	73	8.0	251	27.5	482	52.9	24	2.6	912
8	70	8.2	21	2.4	52	6.1	200	23.3	488	56.9	27	3.1	858
Total	569	8.7	128	2.0	186	2.8	1,689	25.8	3,911	59.7	67	1.0	6,550

Note. Row percentages may not sum to 100 due to rounding.

^a Includes Alaska Natives, Pacific Islanders, and all other groups not classified as African American, American Indian, Asian, Hispanic, or White.

Table A.5 Demographic Characteristics of the Sample, by Grade and Median Family Income

Grade	Median Family Income Level						Total N
	Low		Middle		High		
	N	%	N	%	N	%	
2	545	51.2	253	23.8	266	25.0	1,064
3	381	39.5	336	34.8	248	25.7	965
4	626	61.0	115	11.2	285	27.8	1,026
5	563	64.9	167	19.3	137	15.8	867
6	475	55.4	218	25.4	165	19.2	858
7	384	42.1	177	19.4	351	38.5	912
8	279	32.5	267	31.1	312	36.4	858
Total	3,253	49.7	1,533	23.4	1,764	26.9	6,550

Note. Row percentages may not sum to 100 due to rounding.

Table A.6 Descriptive and Reliability Statistics, by Grade

Grade	Mean ^a	SD ^b	SEM ^c	r ^d	Split-Half ^d	Alpha ^d
2	23.2	12.3	3.78	.86	.89	.87
3	20.0	10.2	3.64	.81	.83	.80
4	26.2	11.4	4.27	.80	.84	.81
5	16.2	10.8	3.54	.84	.89	.84
6	21.2	11.8	3.66	.86	.88	.85
7	20.9	12.2	3.55	.88	.88	.87
8	16.7	10.7	3.20	.86	.88	.85

^aWeighted average.

^bPooled standard deviation.

^cThe SEM for each probe was calculated based on the average correlation coefficient and the actual standard deviation of the raw score for the probe. The average SEM for the grade was calculated by averaging the squared SEMs for each probe and obtaining the square root of the result.

^dThe average reliability coefficients were calculated using Fisher's z transformation.



Appendix B

Administration Directions for Grades 2–6 and Grades 7–8

Grades 2–6

Administration Directions

Before you begin, turn off any cell phones, beepers, or watch alarms. For appropriate grades, make sure all students have their cell phones turned off.

The instructions are carefully worded with simple, age-appropriate language. What you say to the student is in **bold print**.

Say to the students:

We're going to take an 8-minute math test.

Read the problems carefully and work each problem in the order presented. Do not skip around.

If you do not know how to work a problem, mark it with an X and move on. Once you have tried all of the problems in order, you may go back to the beginning of the worksheet and try to complete the problems you marked.

Write the answers to the problems in the blanks. For multiple choice questions, place the letter (A, B, or C) of the correct answer in the blank.

You do not have to show your work, but you may if that is helpful for you in working the problems.

Keep working until you have completed all of the problems or I tell you to stop.

Do you have any questions?

Answer any questions the students may have, hand the students their probes, and say:

Here are your tests.

Put your name, your teacher's name, and the date on each page in the space provided.

Do not start working until I tell you to begin.

Allow the students time to write their information on the probe.

Begin.

If a student asks a question or requests clarification, redirect him or her to the probe and say:

Read the directions again, and work the problem the best you can.

If you still do not understand the problem or are unable to work it, you may move to the next question.

When the appropriate time has elapsed (8 minutes for grades 2–6), say:

Stop and put down your pencil.

If a student(s) continues to work, re-state:

Stop working now and put down your pencil.

At this time, the examiner should collect the probe and proceed to scoring.

Grades 7–8

Administration Directions

Before you begin, turn off any cell phones, beepers, or watch alarms. For appropriate grades, make sure all students have their cell phones turned off.

The instructions are carefully worded with simple, age-appropriate language. What you say to the student is in **bold print**.

Say to the students:

We're going to take a 10-minute math test.

Read the problems carefully and work each problem in the order presented. Do not skip around.

If you do not know how to work a problem, mark it with an X and move on. Once you have tried all of the problems in order, you may go back to the beginning of the worksheet and try to complete the problems you marked.

Write the answers to the problems in the blanks. For multiple choice questions, place the letter (A, B, or C) of the correct answer in the blank.

You do not have to show your work, but you may if that is helpful for you in working the problems.

Keep working until you have completed all of the problems or I tell you to stop.

Do you have any questions?

Answer any questions the students may have, hand the students their probes, and say:

Here are your tests.

Put your name, your teacher's name, and the date on each page in the space provided.

Do not start working until I tell you to begin.

Allow the students time to write their information on the probe.

Begin.

If a student asks a question or requests clarification, redirect him or her to the probe and say:

Read the directions again, and work the problem the best you can.

If you still do not understand the problem or are unable to work it, you may move to the next question.

When the appropriate time has elapsed (10 minutes for grades 7–8), say:

Stop and put down your pencil.

If a student(s) continues to work, re-state:

Stop working now and put down your pencil.

At this time, the examiner should collect the probe and proceed to scoring.



Appendix C

Further Scoring Guidance

After Testing

Although the maximum point value per probe varies modestly across grades, each probe has the exact same point value within a grade. Table C.1 presents the total point value per probe by grade.

Table C.1 Total Available Points by Grade

Grade	Maximum Points
2	45
3	46
4	49
5	51
6	46
7	49
8	42

Using the provided Answer Key, probes should be scored as soon after completion as possible. Generally speaking, scoring M–CAP probes is straightforward. A student’s answer for a particular problem is compared to the correct answer on the Answer Key. Multiple acceptable answers may exist for some items. These alternate answers are provided for those particular problems on the Answer Key.

Figure C.1 presents an example of a completed Answer Key.

Student: *First Last* Teacher: *Ms. Name* Date: *9-5-10*

Reminder: There is **no** partial credit when scoring. The answer must be correct **in its entirety** to obtain the correct score value. If any part of a multi-part question is incorrect, the score is zero.

Grade 2, Probe 4 Answer Key

Item No.	Answer	Correct	Incorrect
1.	9	1	0
2.	401, 957	1	0
3.	6	1	0
4.	14	1	0
5.	56	1	0
6.	7	1	0
7.	–	1	0
8.	28, 32	2	0
9.	45, 50	1	0
10.	24	1	0
11.	95	1	0
12.	610, 629	1	0
13.	5, 45	1	0
14.	0, 15, 40	2	0
15.	+	1	0
16.	C	3	0
17.	$\frac{9}{25}$	2	0
18.	5, 6, 4	2	0
19.	0.85	2	0
20.	less	1	0
21.	E, H, C	2	0
22.	1, 8, 6	1	0
23.	7	2	0
24.	697	1	0
25.	8	2	0
26.	5, 30	2	0
27.	4	3	0
28.	$\frac{11}{15}$	2	0
29.	710	3	0
Total		16	

Figure C.1 Example of a Completed Answer Key

Students may on occasion provide answers that are not easily identifiable as correct or incorrect. In *any* test that requires students to produce written answers, rather than relying on selection-type responses (i.e., multiple choice), there will be some ambiguous answers. In these instances, examiners should understand the common points of potential confusion observed in student responses as provided in this Appendix and then use their professional judgment to determine if the response can be counted as correct. Do not spend an excessive amount of time trying to figure out what the student intended with his or her answer. It should be noted that performance on a single problem should not impact decisions significantly.

Scoring Rules

First and foremost, to be counted as correct, the full answer to the problem must be correct, even when there are multiple answers to the item. This Appendix provides examples of correct and incorrect answers. To determine credit use the following guidelines:

- Does the student's response match the answer (or alternate answers) provided on the Answer Key?
- If the problem has multiple parts, are all parts answered correctly? If not, no credit is given.
- Is the format of the answer correct? If the task is a money task, did the student present the answer properly? For example, was a decimal placed properly between whole dollars and change (e.g., \$4.40 versus \$44.0)?
- Does the answer reflect an understanding of the task type? For example, if the target task is to place unordered numbers (e.g., $\frac{1}{2}$, $\frac{7}{8}$, $\frac{1}{3}$, $\frac{4}{5}$) in sequence from least to greatest (e.g., $\frac{1}{3}$, $\frac{1}{2}$, $\frac{4}{5}$, $\frac{7}{8}$), the answer must reflect the proper sequence. If the student takes additional steps, such as converting a mixed number to a decimal (e.g., .33, $\frac{1}{2}$, $\frac{4}{5}$, $\frac{7}{8}$), the order must still be correct. If there is an error in conversion (e.g., .43, $\frac{1}{2}$, $\frac{4}{5}$, $\frac{7}{8}$), even if the result is in the correct order, the response is incorrect.

This Appendix presents a set of scoring rules and examples by type (error or task) and the grades where these types of errors are most likely to be seen, followed by examples of correct, incorrect, and ambiguous responses. It also provides examples for incomplete responses, flipped or reversed numbers, and over-corrections. Where appropriate, additional explanatory text is provided. The examples in this Appendix *are not exhaustive*; they are representational of responses by students in the national field testing sample. They are provided here to help examiners use their own professional judgment when assessing the correctness of answers that deviate from the correct responses identified on the Answer Key.

Error Type: Mistake in Multi-Part Question

Grades Affected: All

For an item to be considered correct, the entire answer must be correct. As previously stated, there is no partial-credit scoring. If an item has three parts and the student answers one or two parts incorrectly, the entire item is scored as incorrect. Figures C.2 through C.8 present multi-part questions from each grade covered by M-CAP that are incorrect due to incorrect or missing information. If any part of a multi-part item answer is missing or incorrect, the entire item is incorrect.

<p>18 Fill in the blanks.</p> <p>✗ 912 = <u>9</u>[✗] hundreds <u>10</u>[✗] tens <u>2</u>[✓] ones</p>	<p>22 Fill in the blanks.</p> <p>✓ 428 = <u>4</u>[✓] hundreds <u>2</u>[✓] tens <u>8</u>[✓] ones</p>
--	---

Figure C.2 Grade 2 Multi-Part Item (Item 18 incorrect, Item 22 correct)

<p>3 Complete the sequence.</p> <p>✗ 30, 36, 42, <u>48</u>[✓], <u>56</u>[✗]</p>	<p>6 Write the correct number in each blank.</p> <p>✓ 935 = <u>9</u> hundreds <u>3</u> tens <u>5</u> ones</p>
--	--

Figure C.3 Grade 3 Multi-Part Item (Item 3 incorrect, Item 6 correct)

<p>11 Use the table to answer the question.</p> <p>✗</p> <table border="1" style="margin-left: 40px;"> <thead> <tr> <th>Activity</th> <th>Time Taken by Dean</th> </tr> </thead> <tbody> <tr> <td>Sprinting</td> <td>35 minutes</td> </tr> <tr> <td>Cycling</td> <td>40 minutes</td> </tr> <tr> <td>Swimming</td> <td>10 minutes</td> </tr> <tr> <td>Running</td> <td>30 minutes</td> </tr> </tbody> </table> <p>What is the total time taken by Dean for his cross-country marathon practice?</p> <p><u>2</u>[✗] hour <u>55</u>[✓] minutes</p>	Activity	Time Taken by Dean	Sprinting	35 minutes	Cycling	40 minutes	Swimming	10 minutes	Running	30 minutes	<p>11 Use the table to answer the question.</p> <p>✓</p> <table border="1" style="margin-left: 40px;"> <thead> <tr> <th>Activity</th> <th>Time Taken by Dean</th> </tr> </thead> <tbody> <tr> <td>Sprinting</td> <td>35 minutes</td> </tr> <tr> <td>Cycling</td> <td>40 minutes</td> </tr> <tr> <td>Swimming</td> <td>10 minutes</td> </tr> <tr> <td>Running</td> <td>30 minutes</td> </tr> </tbody> </table> <p>What is the total time taken by Dean for his cross-country marathon practice?</p> <p><u>1</u> hour <u>55</u> minutes</p>	Activity	Time Taken by Dean	Sprinting	35 minutes	Cycling	40 minutes	Swimming	10 minutes	Running	30 minutes
Activity	Time Taken by Dean																				
Sprinting	35 minutes																				
Cycling	40 minutes																				
Swimming	10 minutes																				
Running	30 minutes																				
Activity	Time Taken by Dean																				
Sprinting	35 minutes																				
Cycling	40 minutes																				
Swimming	10 minutes																				
Running	30 minutes																				

Figure C.4 Grade 4 Multi-Part Item (Item 11 incorrect, Item 11 correct)

- 2** 9,309,552
- ✗ Which digit is in the millions place?
- 3 ✗
- Which digit is in the thousands place?
- 9 ✓

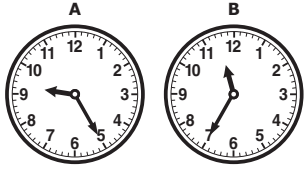
- 3** Maria started cycling at time A. She stopped cycling at time B.
- ✓
- 
- Maria cycled for a total of 2 hours and 10 minutes.

Figure C.5 Grade 5 Multi-Part Item (Item 2 incorrect, Item 3 correct)

- 17** Write the answer in each blank.
- ✗ Note: 16 oz = 1 lb
- 98 oz = 5 ✗ lb 2 ✓ oz

- 20** Write the numbers from least to greatest.
- ✓
- $\frac{11}{13}, 3\frac{9}{12}, \frac{31}{12}, 3, \frac{12}{15}$
- $\frac{12}{15} < \frac{11}{13} < \frac{31}{12} < 3 < 3\frac{9}{12}$

Figure C.6 Grade 6 Multi-Part Item (Item 17 incorrect, Item 20 correct)

- 1** Write these numbers in increasing order.
- ✗
- 0.1, 1.7, $\frac{-3}{2}$, -3.5, 2
- $\frac{-3}{2}$ ✗ -3.5 ✗ 0.1 ✓ 1.7 ✓ 2 ✓

- 6** Write <, >, or = in each blank.
- ✓
- $\frac{10}{11}$ > $\frac{19}{21}$ > $\frac{10}{12}$

Figure C.7 Grade 7 Multi-Part Item (Item 1 incorrect, Item 6 correct)

- 9** Write these numbers in increasing order.
- ✗
- 0.4, $\frac{-7}{2}$, 2.5, $\frac{3}{5}$, -1.4
- $\frac{-7}{2}$ ✓ 0.4 ✗ -1.4 ✗ $\frac{3}{5}$ ✓ 2.5 ✓

- 10** On a blueprint, the dimensions of a park are 25 centimeters (cm) by 10 cm. The blueprint uses a scale of 1 cm = 7 meters (m). What are the actual dimensions of the park?
- ✓
- 175 m by 70 m

Figure C.8 Grade 8 Multi-Part Item (Item 9 incorrect, Item 10 correct)

Error Type: Incomplete Problems

Grades Affected: All

Because there is no partial-credit scoring, an incomplete problem is incorrect. In the case of a multi-part item, if the student leaves any part of the problem incomplete or unanswered, the item is incorrect.

Figure C.9 shows an example of an incomplete problem from grade 2. Although the student has shown the work, and the work indicates that the student was on the right track for a correct answer, the answer itself is not provided; therefore, no credit (0) is given.

- 11** How much money is pictured below?

✗



$$\begin{array}{rcl} \text{_____} \text{¢} & 4 \text{ nickels} & = 20 \\ & 4 \text{ dimes} & = 20 \\ & 8? & \underline{40} \end{array}$$

Figure C.9 Incomplete Problem

Figure C.10 also presents a grade 2 item. In this case, there are three target answers. Although the student has correctly responded with two of the three targets, the answer is incomplete and no credit is given. Note that had the student filled in the last blank with an incorrect answer, the item would still be considered incorrect.

- 18** Fill in the blanks.

✗

973 = _____ hundreds 7 tens 3 ones

Figure C.10 Incomplete Multi-Part Problem

Error Type: Crossed-Out Problem

Grades Affected: All

If a student shows his or her work, but then crossed or X-ed out the problem *without* placing the answer in the blank, the item is incorrect and no points are given. If the student has crossed out the problem, but then returned to the item and placed an answer in the blank, score the item based on whether or not the answer placed in the blank is correct. Figures C.11 through C.13 show examples of each of these possibilities.

- 9 Roberto weighs 120 pounds (lb) and his father weighs 205 lb. How much less is Roberto's weight than his father's weight?

_____ lb

$$\begin{array}{r} \cancel{120} \\ \cancel{205} \\ \hline \end{array} \quad \begin{array}{r} \cancel{205} \\ \cancel{120} \\ \hline 85 \end{array}$$

Figure C.11 Crossed-out Work with No Answer Provided

- 9 Kevin weighs 57 kilograms (kg) and his father weighs 85 kg. How much less is Kevin's weight than his father's weight?

_____ kg

$$\begin{array}{r} \cancel{57} \\ \cancel{85} \\ \hline \end{array} \quad \begin{array}{r} \cancel{85} \\ \cancel{57} \\ \hline 33 \end{array}$$

Figure C.12 Crossed-out Work with Incorrect Answer Provided

- 9 Kevin weighs 115 pounds (lb) and his father weighs 198 lb. How much less is Kevin's weight than his father's weight?

_____ lb

$$\begin{array}{r} \cancel{115} \\ \cancel{198} \\ \hline 7 \end{array} \quad \begin{array}{r} \cancel{198} \\ \cancel{115} \\ \hline 83 \end{array}$$

Figure C.13 Crossed-out Work with Correct Answer Provided

Error Type: Illegible, Reversed, or Rotated Numbers

Grades Affected: All

When students write answers that have illegible, reversed, or rotated numbers, it is important to keep in mind the intent of using M-CAP probes: to determine a student's understanding of the task and progress throughout the school year. Problems with legibility are common, particularly with the younger grades, and students identified as having specific learning challenges may have issues with reversing numbers and letters. Figures C.14 through C.17 provide examples.

If the response is hard to read, but can be determined, score the answer as correct.

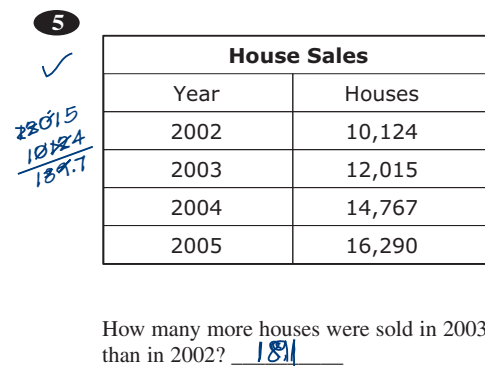


Figure C.14 Difficult-to-Read Response

If the response is too illegible to determine with confidence, score as incorrect.

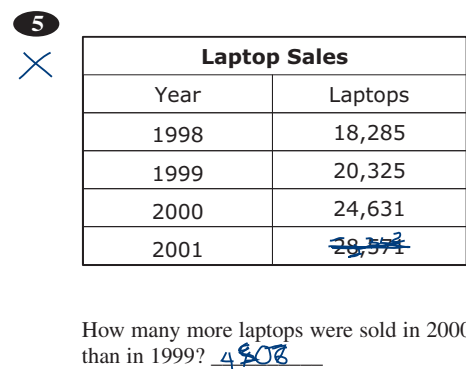


Figure C.15 Illegible Response

If the response is reversed, but it is obvious what digit the student intended, score as correct.

5
✓

Laptop Sales	
Year	Laptops
1998	15,485
1999	<u>16,325</u>
2000	18,831
2001	<u>20,171</u>

How many more laptops were sold in 2001 than in 2000? 1840

Figure C.16 Reversed Numbers

If the response is rotated, and it cannot be easily determined what digit was intended, score as incorrect.

5
✗

Computer Sales	
Year	Computers
2000	25,891
2001	33,539
2002	28,102
2003	37,500

How many more computers were sold in 2003 than in 2000?

87500
25861

11906

Figure C.17 Rotated Numbers

Error Type: Improper Use or Placement of Mathematical Symbols

Grades Affected: All

The proper placement and use of mathematical symbols is a vital part of determining the correctness of an answer. Generally, M–CAP probes provide the appropriate symbols on money, measurement, and time items. If a student also adds a symbol, and does so correctly, the addition will not affect the scoring of that item; however, if a student adds the symbol incorrectly or adds the wrong symbol, the item is considered incorrect.

Examples of correct, incorrect, and variant answers for money, measurement, and time tasks are presented below.

Money Tasks

Figure C.18 presents a typical money task found in the earlier grades.

- 5** How much money is pictured below?



\$ _____

Figure C.18 Grade 3, Item 5

This item presents a picture of a dollar bill and six coins, then asks the student to provide the amount of money shown in the art, and the answer blank provides the dollar (\$) symbol. The correct answer is 1.35. The elements necessary to make the answer correct are the proper number amount and the inclusion and proper placement of the decimal point.

The correct expression of this answer is shown in Figure C.19 on page 44.

- 5 How much money is pictured below?



\$ 1.35

Figure C.19 Grade 3, Item 5, Correct Answer

Had the student added the dollar symbol before the answer, as in Figure C.20, the answer would still be correct.

- 5 How much money is pictured below?



\$ \$1.35

Figure C.20 Grade 3, Item 5, Alternate Correct Answer

Incorrect uses of a symbol in this item type include, but are not limited to, incorrect placement of the dollar sign, using a colon rather than a decimal between the dollar and cents amount, and not including the decimal in favor of using a cent symbol. These are considered incorrect because they indicate a lack of understanding regarding how to present money correctly. See Figures C.21 through C.23 for examples.

- 5 How much money is pictured below?

✗



\$ 1.35¢

Figure C.21 Grade 3, Item 5, Incorrect Dollar Symbol Use

- 5 How much money is pictured below?

✗



\$ 1:35

Figure C.22 Grade 3, Item 5, Improper Use of Colon for Decimal

- 5 How much money is pictured below?

✗



\$ 1¢35

Figure C.23 Grade 3, Item 5, Improper Use of Cent Symbol, Missing Decimal

Corollary to this, the improper inclusion of a symbol can also make a potentially correct answer incorrect. For example, grade 3, Item 14 asks the student to name the coin that is most likely to be pulled at random from a pocket. In the example shown in Figure C.24, the correct answer is “penny.” If a student responds with a numeric value (e.g., 0.01) or any term other than the name of the coin (e.g., cent, 1 cent), the answer is incorrect. See Figure C.25 for another example of an incorrect response.

- ✓ **14** Kayla has 66 pennies, 57 dimes, and 61 nickels in her piggy bank. If she pulls out one coin without looking, what type of coin is she most likely to pull out?

penny

Figure C.24 Grade 3, Item 14, Correct

- ✗ **14** Jessica has 49 pennies, 60 dimes, and 56 nickels in her piggy bank. If she pulls out one coin without looking, what type of coin is she most likely to pull out?

0.10¢

Figure C.25 Grade 3, Item 14, Incorrect

Measurement Tasks

Figure C.26 presents a typical measurement task found in grade 5.

- 25** A house has 60 steps between the first floor and the third floor. Each step is 4 inches (in.) high. What is the distance in feet (ft) between the first floor and the third floor?

Note: 12 in. = 1 ft

_____ ft

Figure C.26 Grade 5, Item 25

As can be seen, the answer blank includes the abbreviation “ft” for “feet.” If a student also includes “ft” or “feet” in his or her answer and it is in the proper placement (after the number), the inclusion does not impact scoring. If, however, the student includes “ft” or “feet” in the wrong location (Figure C.27) or includes the wrong unit of measurement (Figure C.28), the answer is incorrect.

- 25** A house has 60 steps between the first floor and the third floor. Each step is 4 inches (in.) high. What is the distance in feet (ft) between the first floor and the third floor?

Note: 12 in. = 1 ft

ft 20 ft

Figure C.27 Grade 5, Item 25, Incorrect Placement of Unit of Measurement

- 25** A house has 60 steps between the first floor and the third floor. Each step is 4 inches (in.) high. What is the distance in feet (ft) between the first floor and the third floor?

Note: 12 in. = 1 ft

20 cms ft

Figure C.28 Grade 5, Item 25, Incorrect Unit of Measurement Included

Clock-Reading Tasks

Grades 2, 3, and 5 all have clock-reading tasks. Figure C.29 presents a typical clock-reading task found in grade 2.

- 26** Write the time.



_____ : _____

Figure C.29 Grade 2, Item 26

Although, as with money tasks, the M-CAP provides the proper symbol in the answer blank, some students may write the complete answer (hour and minutes) in one of the blanks. In doing so, these students may or may not use the proper symbol to separate the hour and minutes. Figures C.30 and C.31 show instances of a correct and incorrect answer, respectively.

- 26** Write the time.



1:30 : _____

Figure C.30 Grade 2, Item 26, Correct Answer in One Blank

- 26** Write the time.



1,30 : _____

Figure C.31 Grade 2, Item 26, Incorrect Answer in One Blank

Error Type: Estimating, Rounding, Reducing, and Converting

Grades Affected: 4, 5, 6, 7, 8

Estimating and Rounding

Starting in grade 4, students are asked to estimate answers. This estimate may be to the nearest 10s place, 100s place, and so on. Where possible, each item was designed so the answer is the same whether the student rounds before or after completing the problem. Where this was not possible, and two answers can be considered correct depending on when rounding occurred, 2 answers are provided on the Answer Key. Answers should not be rounded unless the item expressly requests that action. Figure C.32 shows an example of an item with two possible correct answers.

- 22** The number of lilies in a nursery is 333.
 ✓ The number of roses is 1343. Estimate (to the nearest hundreds place) the total number of both lilies and roses in the nursery.

<u>1700</u>	$\begin{array}{r} 300 \\ 1300 \\ \hline 1600 \end{array}$	$\begin{array}{r} 333 \\ 1343 \\ \hline 1676 \\ 1700 \end{array}$
-------------	---	---

Figure C.32 Grade 7, Item 22

Reducing Fractions

Unless the item expressly asks students to reduce a fraction, it is not required for a correct answer. However, because many students reduce fractions without being asked, as long as the reduction is correct, the answer is correct. For this reason, where reduction is possible the Answer Key provides both the initial fraction produced by working the problem and the lowest reduction possible of that fraction. Either answer is correct. Every possible reduction between the original fraction and the final reduction is not provided, and are not considered correct. If the student chooses to reduce, the reduction must be the lowest possible reduction. Figures C.33 through C.35 provide two correct and one incorrect reduction examples.

- 23 Write the answer in the blank.



At a carnival, Jesse throws a ring around one of the toys shown above. What is the chance that the ring will fall around a donkey?

$$\frac{2}{4}$$

Figure C.33 Grade 4, Item 23 Correct, Not Reduced

- 23 Write the answer in the blank.



At a carnival, Jesse throws a ring around one of the toys shown above. What is the chance that the ring will fall around a donkey?

$$\frac{1}{2}$$

Figure C.34 Grade 4, Item 23 Correct, Reduced

- 23 Write the answer in the blank.



At a carnival, Jesse throws a ring around one of the toys shown above. What is the chance that the ring will fall around a donkey?

$$\frac{2}{4} \quad \frac{2}{2}$$

Figure C.35 Grade 4, Item 23 Incorrect, Reduced

Converting Mixed Numbers to Decimals

In grades 6 through 8, tasks are included that ask the students to sequence numbers that include mixed numbers and decimals, either from greatest to least or least to greatest. The expectation is that students take the provided numbers and correctly sequence them in the number format presented. Figure C.36 shows a properly sequenced task of this type.

- 20 Write the numbers from least to greatest.

✓

$$\frac{3}{8}, 2\frac{3}{9}, \frac{13}{3}, 2, \frac{3}{7}$$

$$\frac{3}{8} < \frac{3}{7} < 2 < 2\frac{3}{9} < \frac{13}{3}$$

Figure C.36 Grade 6, Item 20 Completed Correctly

Some students may decide to reduce the fractions and to convert the resulting mixed numbers to decimals. If the reduction and conversion are done correctly and then sequenced correctly, the answer is correct (see Figure C.37). If the student makes an error in reduction or conversion, the answer is incorrect, whether not the final sequencing is correct based on the incorrect conversion (see Figure C.38).

- 20 Write the numbers from least to greatest.

✓

$$\frac{3}{8}, 2\frac{3}{9}, \frac{13}{3}, 2, \frac{3}{7}$$

$$\frac{3}{8} < \frac{3}{7} < 2 < 2\frac{3}{9} < 4\frac{1}{3}$$

Figure C.37 Grade 6, Item 20 Correctly Converted and Completed

- 20 Write the numbers from least to greatest.

✗

$$\frac{3}{8}, 2\frac{3}{9}, \frac{13}{3}, 2, \frac{3}{7}$$

$$\frac{3}{8} < \frac{3}{7} < 2 < 2\frac{3}{9} < 4\frac{1}{2}$$

Figure C.38 Grade 6, Item 20 Incorrectly Converted and Completed

In grades 6 through 8, students are also asked to complete number sequences that include mixed numbers and decimals. The same rule applies to this item type as to the items in Figures C.36–C.38. Figure C.39 presents a correctly completed sequence without conversion; Figure C.40 presents a correctly completed sequence with correct conversion; Figure C.41 presents an incorrectly completed sequence with incorrect conversion.

25 Complete the sequence.



$$\frac{1}{4}, \frac{5}{4}, \frac{9}{4}, \frac{13}{4}, \underline{\frac{17}{4}}$$

Figure C.39 Grade 7, Item 25 Completed Correctly

25 Complete the sequence.



$$\frac{1}{4}, \frac{5}{4}, \frac{9}{4}, \frac{13}{4}, \underline{4\frac{1}{4}}$$

Figure C.40 Grade 7, Item 25 Completed Correctly with Correct Conversion

25 Complete the sequence.



$$\frac{1}{4}, \frac{5}{4}, \frac{9}{4}, \frac{13}{4}, \underline{4\frac{1}{3}}$$

Figure C.41 Grade 7, Item 25 Completed Incorrectly with Incorrect Conversion



References

- American Educational Research Association, American Psychological Association, & National Council on Measurement Education. (1999). *Standards for educational and psychological testing*. Washington, DC: Author.
- Fuchs, L. S., & Fuchs, D. (1987). *Effects of curriculum-based measurement procedures in spelling and math*. Unpublished manuscript available from L. S. Fuchs, Box 328, Peabody College, Vanderbilt University, Nashville, TN 37203.
- Fuchs, L. S., & Shinn, M. R. (1989). Writing CBM IEP Objectives. In M. R. Shinn (Ed.), *Curriculum-based measurement: Assessing special children*. (pp. 132–154). NY: Guilford.
- Fuchs, L. S., Fuchs, D., & Hamlett, C. L. (1990). Curriculum-based measurement: A standardized long-term goal approach to monitoring student progress. *Academic Therapy*, 25(5), 615–632.
- Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Stecker, P. M. (1991). Effects of curriculum-based measurement and consultation on teacher planning and student achievement in mathematics operations. *American Educational Research Journal*, 28, 617–641.
- Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Whinnery, K. W. (1991). Effects of goal line feedback on level, slope, and stability of performance within curriculum-based measurement. *Learning Disabilities Research and Practice*, 6, 3–12.
- Fuchs, L. S., Fuchs, D., & Zumeta, R. O. (2008). A curricular sampling approach to progress monitoring: Mathematics concepts and applications. *Assessment for Effective Intervention*, 33, 225–233.
- Mathematics Learning Study Committee & National Research Council. (2001). *Adding it up: Helping children learn mathematics*. Washington, DC: National Academy Press.
- Mirkin, P. K., Deno, S. L., Fuchs, L. S., Wesson, C., Tindal, G., Marston, D., et al. (1981). *Procedures to develop and monitor progress on IEP goals*. Minneapolis: University of Minnesota Institute for Research on Learning Disabilities.
- National Council of Teachers of Mathematics (2006). *Curriculum focal points for prekindergarten through grade 8 mathematics: A quest for coherence*. Reston, VA: Author.
- National Council of Teachers of Mathematics. (1989). *Principles and standards for school mathematics*. Reston, VA: Author.
- National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: Author.
- National Mathematics Advisory Panel. (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel*. Washington, DC: U.S. Department of Education.

- Shinn, M. R. (2002a). *Progress monitoring strategies for writing individualized goals in general curriculum and more frequent formative evaluation*. Eden Prairie, MN: Edformation, Inc.
- Shinn, M. R. (2002b). *Organizing and implementing a benchmark assessment program*. Eden Prairie, MN: Edformation, Inc.
- Shinn, M. R. (2008). Best practices in curriculum-based measurement and its use in a problem-solving model. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology* (5th ed., pp. 243–262). Bethesda, MD: National Association of School Psychologists.
- Shinn, M. R., & Shinn, M. M. (2000). Writing and evaluating IEP Goals and making appropriate revisions to ensure participation and progress in general curriculum. In C. F. Telzrow and M. Tankersley (Eds.), *IDEA Amendments of 1997: Practice guidelines for school-based teams*. (pp. 351–381). Bethesda, MD: National Association of School Psychologists.
- Tindal, G., Germann, G., & Deno, S. (1983). *Descriptive research on the Pine County norms: A compilation of findings*. (Research Report No. 132). Institute for Research on Learning Disabilities.